DOCUMENT RESUME

ED 458 874 IR 021 028

AUTHOR	Hwang, Young Suk; Fisher, William; Vrongistinos, Konstantinos
TITLE	Calibrating a Measure of Motivation in Using Technology.
PUB DATE	2001-04-00
NOTE	17p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001). Contains colored chart with illegible type.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Educational Technology; Grade 6; Intermediate Grades; *Learning Motivation; *Measurement Techniques; *Measures (Individuals); *Student Motivation; Student Surveys

ABSTRACT

This study examines the functioning of an instrument for measuring elementary school students' motivation to learn about technology. The theoretical framework of the study is based on expectancy-value theory, which focuses on cognitive factors in determining achievement behaviors. Participants were 129 sixth grade students. A questionnaire was group administered orally to the students. The 36 items in the questionnaire addressed the following areas thought to bear on a measure of motivation to learn about technology: self-concept of ability, perception of technology, intrinsic causal attribution, extrinsic causal attributions, task-involved motivation, ego-involved motivation, parents' perception, female gender issues, and male gender issues. It is concluded that the attempt to develop an instrument to measure motivation in learning about technology was successful. Several figures and tables presenting survey data are appended. (Contains 17 references.) (MES)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 458 874

Calibrating a Measure of Motivation in Using Technology

Young Suk Hwang, CSUSB

William Fisher, LSU

Konstantinos Vrongistinos, CSUN

A paper presented at the meeting of the

American Educational Research Association

Seattle, Washington

April, 2001

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Y.S. Hwang

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.

Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Correspondence concerning this paper should be directed to the first author at the following address: Young Suk Hwang, Ph.D. California State University, San Bernardino, Department of Educational Psychology and Counseling, 5500 University Parkway, San Bernardino, CA 92407-2397 Email yhwang@csusb.edu Phone: 909-880-5672, Fax: 909-880-7040

IR021028

BEST COPY AVAILABLE

Calibrating a Measure of Motivation in Learning about Technology

American society is increasingly technologically oriented. Consequently, our keeping pace with other industrialized nations requires educators to insure that students who are future entrants into the workforce have the necessary knowledge and skill in learning and using technology/computer. However, research and statistical data demonstrated that American students have fallen off the pace with regard to mathematics and science achievement.

Even within American population, it is apparent that there are groups that are not performing up to their potential. For example, females (Linn & Hyde, 1988) have been shown to achieve less in science and math than white middle class males. Because demographic trends suggest that woman and minorities will constitute larger percentages of the workforce in the future, it is imperative that we understand why females are not achieving up to their potential and are not choosing to pursue careers in technology/computer.

OBJECTIVES

The present study examines the functioning of an instrument for measuring elementary school students' motivation to learn about technology. If this variable can be measured quantitatively, as shown via data fit to a fundamental measurement model (Luce & Tukey, 1964; Krantz, Luce, Suppes, & Tversky, 1971; Michell, 1990, 1997; 1999), further research will examine potential sources of differences in learning and using technology between female and male sixth graders. The theoretical framework of this study is based on Eccles and Wigfield (1995) expectancy-value theory, which focuses on cognitive factors in determining achievement behaviors. More specifically, differences in self-concept of ability, perception of technology,

perception of parental beliefs, and causal attributions (success and failure), are investigated as they relate to achievement.

METHOD

Participants

There were 129 (65 males & 64 females) sixth grade students enrolled in an elementary school in a southern state. The mean age for the students was 10.4.

Instrument

The questionnaire used in this study was revised based on Whang and Hancock (1994) and Eccles' expectancy x value theory. The 36 items in the questionnaire thought to bear on a measure of motivation to learn about technology involve self-concept of ability (6 items), perception of technology (4 items), intrinsic casual attributions (4 items), extrinsic casual attributions (3 items), task-involved motivation (3 items), ego-involved motivation (3 items), parent's perception (7 items), female gender issues (3 items), male gender issues (3 items). These items are constructed using a Likert scale anchored by very true (1) and not true at all (4).

Procedure

The surveys were group administered orally to students in their homerooms, using standardized instructions by a trained graduate student. Participants were told that the purpose of the study was to learn what students think using and learning technology/computer is all about and how they feel about their ability in using and learning technology/computer. Approximately 30 min was needed to complete the questionnaire. Care was taken through the surveys to give no indication of what would be considered the appropriate or "right" answer. If the student did not

understand a question, it was repeated or paraphrased in simpler language, but still no examples or suggested answers were given.

Analysis

Fundamental measurement theory (Luce & Tukey, 1964; Krantz, Luce, Suppes, & Tversky, 1971; Michell, 1990, 1997; 1999) provides a strong program for testing the hypothesis that a given variable is quantitative. According to Mundy (Mundy, 1986, p. 392),

> The hallmark of a meaningless proposition is that its truth-value depends on what scale or coordinate system is employed, whereas meaningful propositions have truth-value independent of the choice of representation, within certain limits.

Fundamental measurement theory's approach to testing the quantitative hypothesis examines the extent to which mathematical propositions concerning the variable will be meaningful or meaningless. In other words, fundamental measurement theory establishes whether a construct's quantitative expression depends upon the particular test or survey questions asked (what might be called the "brand" of instrument) and/or upon the particular sample of persons responding to the questions.

A particularly easy to use and convenient way of testing the quantitative hypothesis is via fit to a probabilistic conjoint (Rasch) measurement model (Rasch, 1960; Perline, Wright, & Wainer, 1979; Wright & Masters, 1982; Andrich, 1988; Fisher & Wright, 1994; Wright, 1999). When data fit one of these models, the axioms of simultaneous conjoint measurement and the meaningfulness criterion are satisfied. The model employed here is the rating scale model

$$nl(p / (1-p)) = b_n - d_i - k_j$$

read as the natural logarithm of the odds (p / (1-p)) that difference between the ability b of person n is greater than the difficulty d of item i at the level k posed by category j. Survey applications typically assume that measures are not affected by any factors other than the properties of the questions asked and the attitudes or abilities of the respondents. Fundamental measurement models do nothing more than check these assumptions for their tenability in the face of actual observations.

Given the sample size of 129 and the 36-item instrument, overall success in the measurement effort will be indicated when the standard deviations of the information-weighted and outlier-sensitive model fit statistics are less than 2.0, and when individual scores function as sufficient statistics (i.e., the pattern of responses across the items for a person, or the pattern of responses across the persons for an item, is reproducible from the score alone) (Smith, 1986, 1998; Wright & Masters, 1982).

RESULTS

Table 1 shows the summary statistics for the student measures and the item calibrations. The data matrix of 129 students times the 36 items contains 4,644 possible observations; 98.8% of these (4,587) are valid. Raw scores range from 67 to 115, within the maximum range of 36 to 144. The students responded to 35.6 questions on average.

Overall model fit appears acceptable for the item calibrations, given that the standard deviations of the standardized fit statistics are considerably less than 2.0.

Several students, however, appear to have provided inconsistent responses, as the maximum information-weighted (infit) and outlier-sensitive (outfit) statistics are 5.7 and 5.2, respectively. Examination of the residuals indicates several highly unexpected responses to items

21 and 10, which ask the students about their parents' opinions concerning girls' need to study technology (21) and concerning whether the respondent hates technology (10).

Modeled reliability, at .62, allows for barely 2 statistically distinct strata (measurement ranges with centers at least three errors apart (Wright & Masters, 1982). The average measurement error of .19 is close to that predicted by Rasch generalizability theory for a survey of 36 four-category items, but the remarkably low measurement standard deviation (SD) of .29 makes for a measurement separation (ratio of the SD to the error) of only 1.29.

Table 2 shows the summary statistics for the rating scale categories. Note that the observed counts in each category range from 749 to 1714, and the step calibrations are not ordered from less to more as the category labels progress from 1 to 4.

The distribution of the measures, averaging .18 logits, is illustrated in Figure 1, showing that the measures are most closely aligned with the calibrations of the items on the average step of the rating scale, which approximates the step from Somewhat True to Not Very True. The average measure is about one error above the center of the item scale (0.0). Were the instrument better targeted, error would be somewhat lower, and reliability, higher.

Figure 2 shows the items in measure order, along with the positions of the category transitions on the number line. The statements at the bottom of the figure are rated least true, and those at the top, most true. Students rate the assertion that their parents do not value technology education important for girls untrue, and they deny learning about technology to prevent getting in trouble. Students also tend to find statements concerning the technical gender superiority of either boys or girls untrue.

At the other end of the continuum, statements that the students find very true involve learning about technology because it is interesting, obtaining good technology course grades because of hard work, expecting good grades in technology classes, and the importance parents place on learning about technology.

Figure 3 shows a box plot of the items' calibration values by the theoretical construct groupings. Items involving task-orientation, ability self-concepts, perceptions of technology, and internal causal attributions are rated most true, with statements concerning ego-involved motivation, external causal attributions, and gender issues rated least true.

Analysis of the measure, error, standardize fit statistic, and point biserial correlation variances (via ANOVA) revealed few relationships that could clearly be interpreted as statistically or substantively significant. The most suggestive findings indicate that the information-weighted fit statistics are elevated for black males and Hispanic females (there were no Hispanic male respondents and only two Hispanic females), meaning that the survey may not validly measure attitudes toward technology for these groups.

CONCLUSIONS and DISCUSSION

Based on these data, it seems that our attempt to develop an instrument to measure motivation in learning about technology was successful, to certain degree. In general, the survey items behaved fairly well as measures of motivation in learning about technology and gender differences. The homogeneity of the measures, as indicated by the low SD, and the high model fit statistics suggest that further research into the construct and the items bearing on it may be warranted. However, eight of the items exhibited low point- biserial correlations. Removing the

four worst-fitting cases and another four individual observations on item 21 reduced the standardized outfit SD from 2.1 to 1.8, and the standardized infit SD from 2.3 to 2.2.

Ideally, each category's probability curve should have its own peak, indicating the point at which it is the most probable response for a given range of measures. The curves in Table 2 indicate that categories 2 and 3 are never the most likely response. Respondents apparently cannot distinguish four separate degrees of truth in these statements. Future analyses of these data should explore various ways of combining categories to linearize the step calibrations. And if the instrument is to be administered to another sample, the response options ought to be expanded from the current four to six, with the category labels modified to more clearly demarcate increasing amounts. For instance, an Agree/Disagree continuum, with "Very Strongly", "Strongly", and "Mildly" as modifiers, might more clearly convey a wider range of possible responses than the current labels' focus on truth.

REFERENCES

Andrich, D. (1988). *Sage University Paper Series on Quantitative Applications in the Social Sciences. Vol. series no. 07-068: Rasch models for measurement.* Beverly Hills, California: Sage Publications.

Eccles, S., and Wigfield, A. (1995). In the mind of the actor: the structure of adolescents' achievement task values and expectance-related beliefs. *Personality and Social Psychology Bulletin, 21,* 215-225.

Fisher, W. P., Jr., & Wright, B. D. (Eds). (1994). Applications of Probabilistic Conjoint Measurement. *International Journal of Educational Research, 21*(6), 557-664.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement. Volume 1: Additive and polynomial representations.* New York: Academic Press.

Linn, M., & Hyde, J. S. (1988). Gender differences in verbal ability: A meta-analysis. *Psychological-Bulletin, 104*(1), 53-69.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new kind of fundamental measurement. *Journal of Mathematical Psychology, 1*(1), 1-27.

Michell, J. (1990). *An Introduction to the Logic of Psychological Measurement.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Michell, J. (1997). Quantitative science and the definition of measurement in psychology. *British Journal of Psychology, 88,* 355-383.

Michell, J. (1999). *Measurement in psychology: A critical history of a methodological concept.* Cambridge: Cambridge University Press.

Mundy, B. (1986). On the general theory of meaningful representation. *Synthese, 67,* 391-437.

Perline, R., Wright, B. D., & Wainer, H. (1979). The Rasch model as additive conjoint measurement. *Applied Psychological Measurement, 3,* 237-255.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests* (reprint, with Foreword and Afterword by B. D. Wright, Chicago: University of Chicago Press, 1980). Copenhagen, Denmark: Danmarks Paedogogiske Institut.

Smith, R. M. (1986). Person fit in the Rasch model. *Educational and Psychological Measurement, 46,* 359-372.

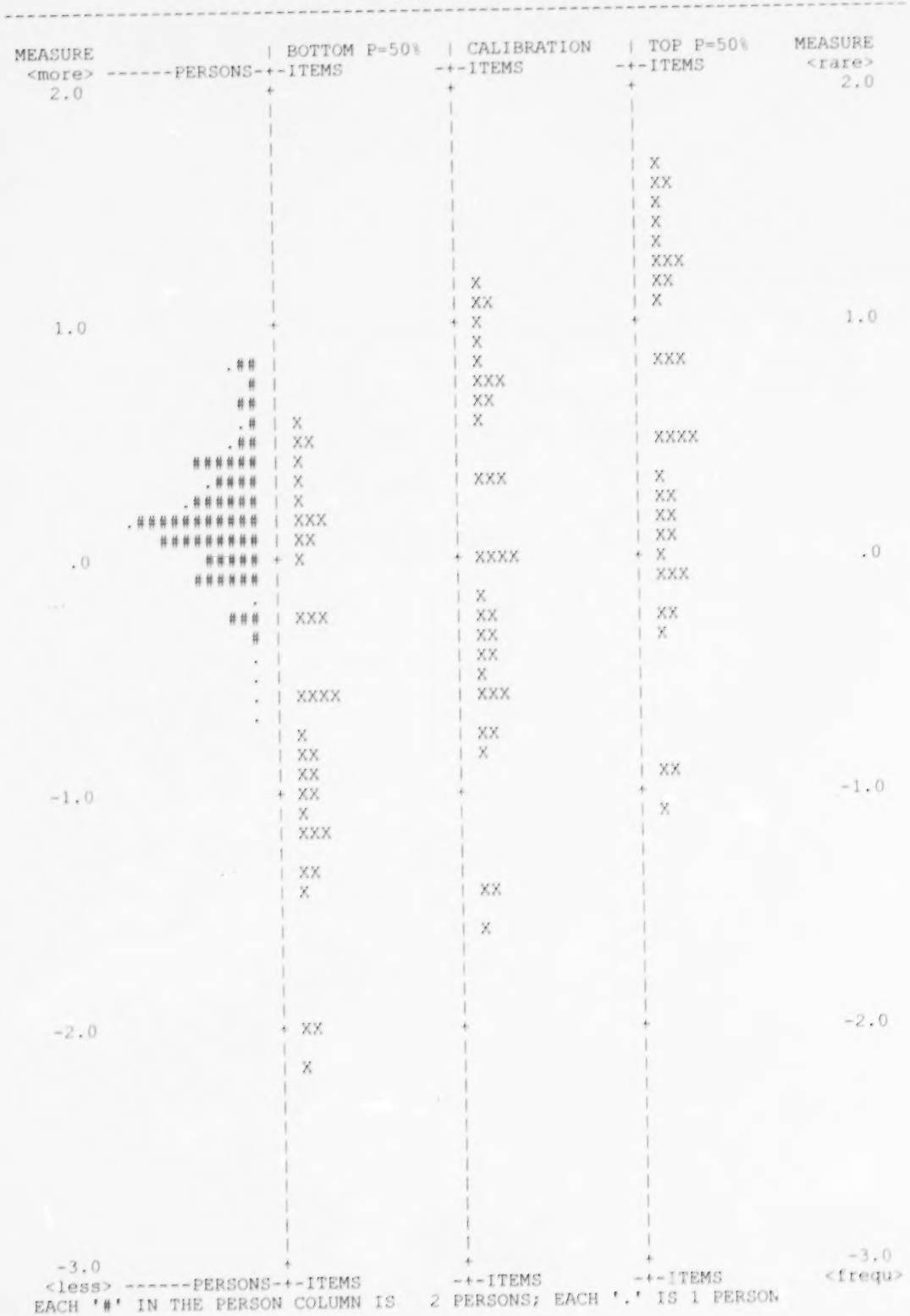
Smith, R. M. (1998). Using item mean squares to evaluate fit to the Rasch model. *Journal of Outcome Measurement, 2*(1), 66-78.

Whang, P. A. & Hancock, G. (1994). Motivation and mathematics achievement: Comparisons between Asian-American and non-Asian students. *Contemporary Educational Psychology, 19*(3), 302-322.

Wright, B. D. (1999). Fundamental measurement for psychology. In S. E. Embretson & S. L. Hershberger (Eds.), *The new rules of measurement: What every educator and psychologist should know.* Hillsdale, NJ: Lawrence Erlbaum Associates.

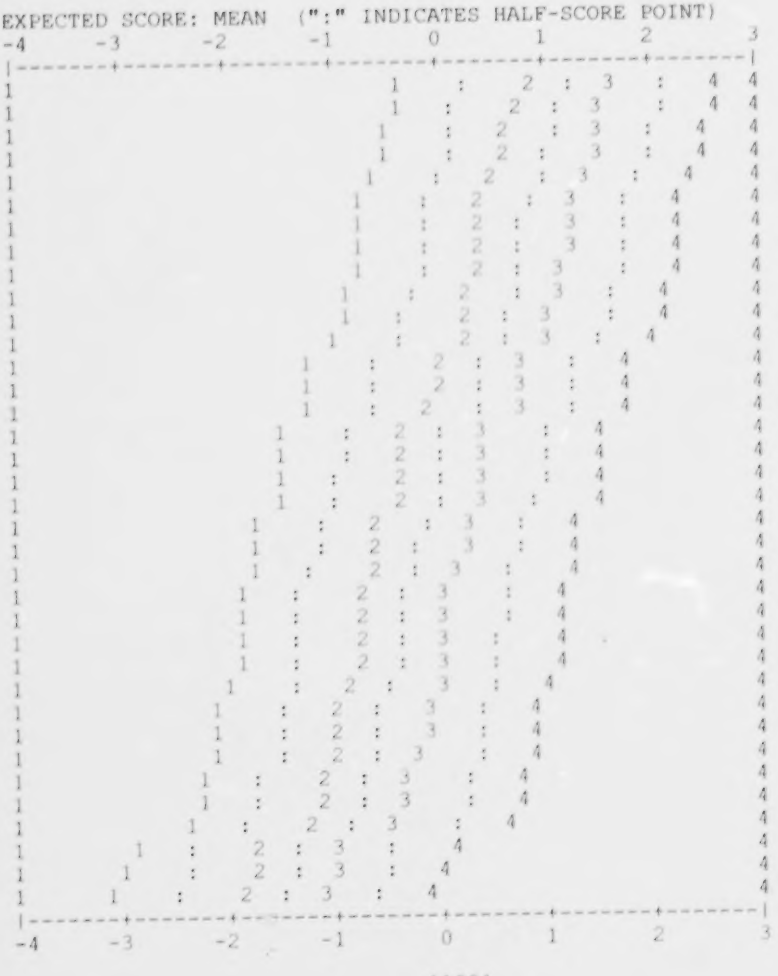
Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis: Rasch measurement.* Chicago: MESA Press.

FIGURE 1
Technical Motivation Variable Map
129 students, 36 items, 4 Categories



BEST COPY AVAILABLE

FIGURE 2
Technical Motivation Expected Responses
129 students, 36 items, 4 Categories

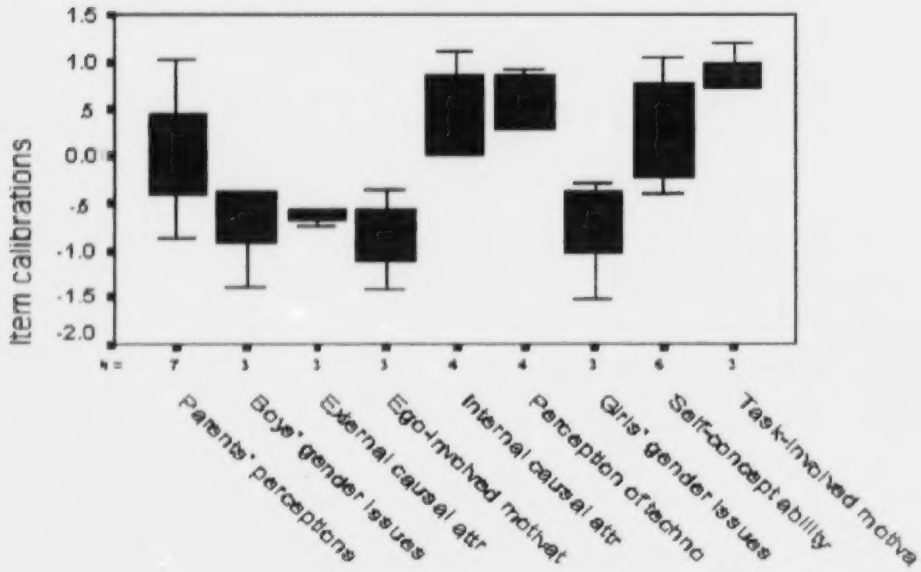

EXPECTED SCORE: MEAN (":" INDICATES HALF-SCORE POINT)
-4 -3 -2 -1 0 1 2 3
NUM ITEM
6 TThe reason I try to learn how
19 IWhen I get a good grade in te
8 SI expect good grades (high pe
7 AMy parents think very importa
3 PTechnology are easy to learn.
25 PUsing technology requires me
26 SI am happy with the grades I
14 TUsing technology helps me stu
23 TTo do well in using technolog
20 SI feel I can do better in tec
4 II am good in using technology
24 AMy parents think that I can d
12 PTechnology projects are time-
22 PTechnology are mainly memoriz
16 AMy parents help me to use tec
35 IWhen I get a bad grade in tec
29 Ibad grade in technology / com
36 AMy parents think that I am on
5 SI work harder in using techno
33 AMy parents think I need to sp
1 SMy classmates are better at u
17 RGirls are better at using com
28 GI try to learn technology tea
34 BBoys are better technology i
18 STechnology projects are diffi
11 BBoys are better at using comp
31 RGirls are better technology
27 EWhen I get a good grade techn
32 Ebad grade in technology / com
30 AMy parents think that using t
13 EI am good in using technology
15 GI try to learn technology is
10 AMy parents think I hate techn
2 BMy parents think learning tec
9 GI try to learn technology get
21 RMy parents think learning tec
NUM ITEM
-4 -3 -2 -1 0 1 2 3
11221
11136382768625
T S M S T
PERSON

BEST COPY AVAILABLE

Attitudes toward Learning about Technology

Calibrations from 129 6th-Grade Students

By item group



3.

BEST COPY AVAILABLE

FIGURE 4
Technical Motivation Principal Components (Standardized Residual) Factor Plot
129 Persons, 36 Items, 4 Categories

Factor 1 explains 4.45 of 36 variance units
Pearson Correlation: -.84

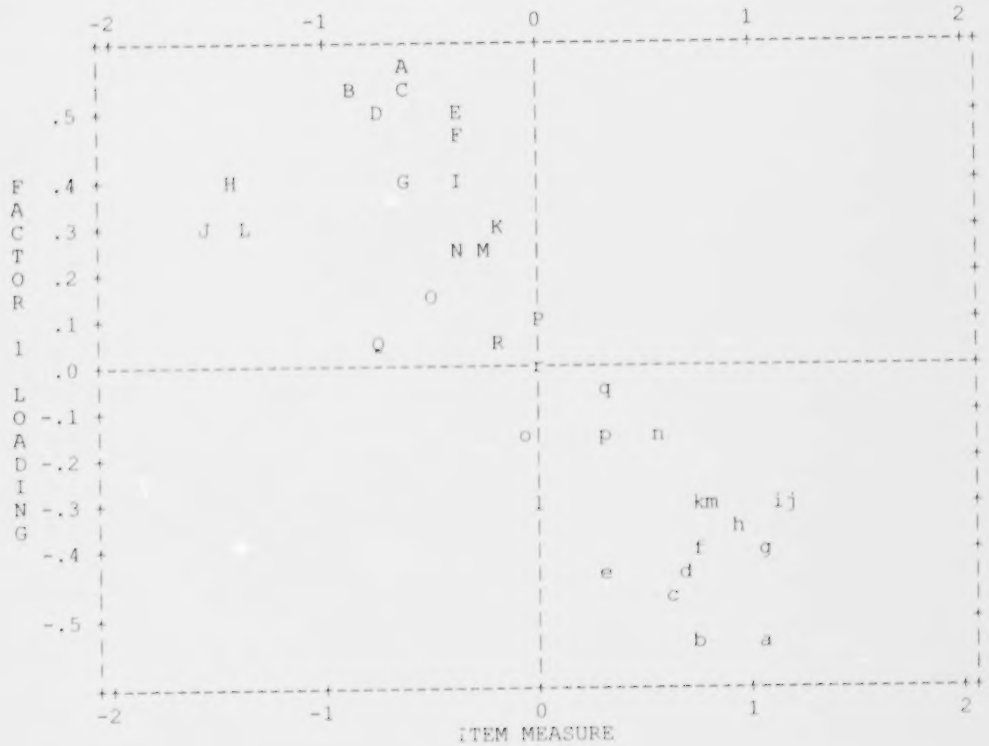


TABLE 1
Technical Motivation Instrument Calibration Summary Statistics
129 students, 36 items, 4 Categories

SUMMARY OF 129 MEASURED PERSONS

	RAW SCORE	COUNT	MEASURE	REAL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	94.0	35.6	.18	.19	1.00	-.4	1.07	-.2
S.D.	9.6	1.6	.29	.03	.53	2.3	.75	2.1
MAX.	115.0	36.0	.87	.29	2.79	5.7	4.52	5.2
MIN.	67.0	21.0	-.70	.17	.33	-4.2	.32	-3.4

REAL RMSE	.19	ADJ.SD	.21	SEPARATION	1.10	PERSON RELIABILITY .55
MODEL RMSE	.18	ADJ.SD	.23	SEPARATION	1.29	PERSON RELIABILITY .62
S.E. OF PERSON MEAN	.03					

VALID RESPONSES: 98.8%

SUMMARY OF 36 MEASURED ITEMS

	RAW SCORE	COUNT	MEASURE	REAL ERROR	INFIT MNSQ	INFIT ZSTD	OUTFIT MNSQ	OUTFIT ZSTD
MEAN	336.7	127.4	.00	.10	1.04	-.1	1.07	.2
S.D.	93.6	1.5	.75	.03	.19	1.4	.23	1.3
MAX	487.0	129.0	1.20	.23	1.64	1.9	2.01	2.6
MIN.	188.0	123.0	-1.54	.08	.67	-4.1	.73	-3.5

REAL RMSE	.11	ADJ.SD	.74	SEPARATION	6.80	ITEM RELIABILITY .98
MODEL RMSE	.10	ADJ.SD	.74	SEPARATION	7.33	ITEM RELIABILITY .98
S.E. OF ITEM MEAN	.13					

BEST COPY AVAILABLE

TABLE 2
Technical Motivation Summary of Measured Steps on the Rating Scale
129 students, 36 items, 4 Categories

SUMMARY OF MEASURED STEPS

CATEGORY LABEL	OBSERVED COUNT	MEASURE AVRGE	MEASURE EXPECT	COHERENCE M->C	COHERENCE C->M	INFIT MNSQ	OUTFIT MNSQ	STEP CALIBRATN	
1	1229	-.44	-.47	71%	18%	1.07	1.27	NONE	Very True
2	895	-.14	-.11	29%	53%	.86	.80	.02	Somewhat True
3	749	.28	.31	23%	59%	.88	.84	.28	Not Very True
4	1714	.76	.74	81%	35%	1.00	1.06	-.30	Not True at All

AVERAGE MEASURE is mean of measures in category.
M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

CATEGORY LABEL	STEP CALIBRATN	STEP S.E.	SCORE-TO-MEASURE AT CAT.	----ZONE----		THURSTONE THRESHOLD	
1	NONE		(-1.52)	-INF	-.95		Very True
2	.02	.04	-.37	-.95	.02	-.55	Somewhat True
3	.28	.04	.42	.02	.96	.06	Not Very True
4	-.30	.04	(1.47)	.96	+INF	.52	Not True at All

BEST COPY AVAILABLE